AD-771 699

THE THEORY AND APPLICATION OF DECISION ANALYSIS

Ronald A. Howard

Stanford University
Stanford, California

1 February 1973

TECHNICAL REPORT

to

ADVANCED RESEARCH PROJECTS AGENCY

on

ARPA Order Number 2449, 1 February 1973
Program Code Number 3D20

for

A Program of Research

on

THE THEORY AND APPLICATION OF DECISION ANALYSIS

Ronald A. Howard, Principal Investigator
(415) 321-2300, Extension 4176

Contract Number: N00014-67-A-0012-0077

Amount of Contract: $50,000.00

Effective Date of Contract: 1 May 1973

Contract Expiration Date: 1 November 1973

Date of Report: 1 December 1973

Contract Period Covered: 5/1/73-11/1/73

Scientific Officer: Director Engineering Psychology Programs Psychological Science Division, Office of Naval Research, Department of the Navy, 800 North Quincy Street, Arlington, Virginia 22217

Name of Contractor: The Board of Trustees of the Leland Stanford Junior University, c/o Office of Research Administrator, Encina Hall, Stanford, California 94305

Form Approved: Budget Bureau - No. 22-R0293

9

TECHNICAL REPORT

RESEARCH PROGRAM PURPOSE

The purpose of this research program is to develop new concepts and procedures for the logical analysis of decisions involving uncertainty, complexity, and dynamic effects.

SUMMARY OF ACCOMPLISHMENTS

The first six months of the research program in decision analysis have seen the completion of one major project and significant progress on three other topics.

The research that is essentially completed concerns the value of information in situations where the successive pieces of information may be obtained sequentially. Results of this research have shown that we must be extremely careful in using conventional value of information calculations in sequential situations. A final report on this research is now being prepared for both distribution and publication*; consequently, we shall not comment further on it at this time.

The three main research projects now in progress are concerned with

1) the economic balancing of assessment and computation in decision analysis,

2) the sensitivity of recommendations to choices made in the modeling process, and

3) the value of retaining flexibility in the setting of decision variables and of having complete as opposed to quantized information.

---

* "The Value of Sequential Information" by Allen Clinton Miller, III, to be published in December 1973.

We shall now describe progress on each of these projects.

RESEARCH PROGRESS

THE ECONOMICS OF ASSESSMENT AND COMPUTATION

The economics of decision analysis addresses the challenging area of problem formulation. The object of the research is to develop formal aids to balance accuracy against computational simplicity. The ultimate result is expected to include a framework general enough to identify most of the decision problems that the professional analyst faces. The framework will structure the new results, identify pertinent work in the literature, and illuminate areas for future work.

The initial state of information is summarized by the decision maker's estimates of the best decision, the mean and variance of the resulting lottery, and the amount that these quantities would change if they were based on a complete analysis. Each possible assessment and computational program is an experiment providing more information about these quantities. The experiments are evaluated using decision theory, with computations simplified using Taylor series approximations.

The research has concentrated on problems with a single decision variable and many normally-distributed state variables. The sequential value of encoding state variables has been computed; the expected gains from Monte Carlo sampling and from decision trees have been evaluated. As a result, the resources available for a project can be divided between computation and assessment. The optimum number of variables to be encoded and the optimum Monte Carlo sample size or tree size can be computed. The

techniques have been applied in an extensive example, based on the actual decision analysis.

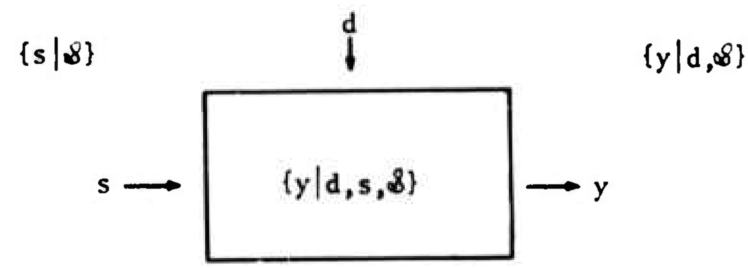
The continuing research will focus on broader decision problems. For example, discrete state variables can make a decision problem unsuitable for analysis based on Taylor series. However, by breaking the problems into blocks, the discontinuities can be removed and the approximate techniques can be applied to each portion.

The research so far has focused on the decisions made between the deterministic and probabilistic phases of the decision analysis. When the cost of evaluating the deterministic model is not large, this is the critical decision point. However, when the cost of evaluating the model is very high, as it would be, for example, if an evaluation of the model requires the deterministic simulation of a national electrical power system over many years, then even a complete deterministic analysis may be a prohibitively expensive luxury. For this class of problems it may be best to do an immediate Monte Carlo simulation and replace the sensitivity analysis with least-squares approximations.

THE ADEQUACY OF MODELS

The value of any decision analysis depends in part on the adequacy of the mathematical models used in that analysis. The object of the present research is to develop a methodology with which the adequacy of the models used in a decision analysis can be quickly and simply determined.

A typical decision situation can be represented as follows:



The outcome variable y is dependent on a decision variable d, which is controlled by the decision maker, and on a state variable s, which is a random variable with probability distribution $\{s|\mathcal{S}\}$. The dependence of y on d and s is embodied in the conditional probability distribution $\{y|d,s,\mathcal{S}\}$. For any value of d, the corresponding probability distribution on y, called the profit lottery, can be found:

$$\{y|d,\mathcal{S}\} = \int_s \{y|d,s,\mathcal{S}\} \cdot \{s|\mathcal{S}\} \quad , \qquad 1)$$

where $\int_s$ denotes a summation over the variable s. The goal of a decision analysis is to identify the value of d that corresponds to the most desirable profit lottery.

In practice, because of the large assessment and computational requirements implied by 1), the profit lottery is almost never calculated in this way. Rather, the decision situation is usually simplified through the use of models:

a) The dependence of y on d and s is modeled as a deterministic function g : $y = g(d,s)$.

b) The probability distribution on s is modeled as a simple (usually discrete) function p : $\{s|\mathcal{S}\} = p(s)$.

Then, with these models, the profit lottery is calculated:

$$\{y|d,\mathcal{S}\} = \int_s \delta(y - g(d,s)) \cdot p(s) \qquad 2)$$

The models are adequate for the purposes of the analysis if, and only if, the profit lottery in 2) is not too different from the profit lottery in 1). However, since the profit lottery in 1) is not known, the discrepancy between the two profit lotteries can only be approximated.

In the present research, methods are being developed to characterize explicitly the error incurred in the modeling assumptions a) and b) and to approximate the discrepancy between the profit lotteries caused by these modeling errors. The error in assumption a) is characterized by a random variable e such that $y = g(d,s) + e$. The error in b) is characterized by the error in the mean and variance of s caused by using $p(s)$ rather than $\{s|\mathcal{S}\}$. The discrepancies in the mean and variance of the two profit lotteries are then shown to be functions of these errors, of moments of s , and of derivatives of g .

When this research is completed, it should be possible to check quickly the adequacy of the models used in any decision analysis.

FLEXIBILITY AND INFORMATION QUANTIZATION

A particularly useful contribution decision analysis has made to the science of decision-making is a logical technique for placing a dollar value on information. We are currently extending these techniques so as to

find the economic value of information systems; that is, we are searching for techniques for determining in the context of a particular problem when information should be collected, how it should be transmitted, and where it should be used. By weighing the economic value versus the cost of various informational systems, the decision maker will have a valuable tool for the design and analysis of systems for the accumulation, transmission, and incorporation of data.

How many times have we heard the phrase, "If I only knew then what I know now." Such a phrase points out the simple fact that the value one places on information is strongly influenced by one's "flexibility"; that is, his ability to adjust his decision strategy in response to reported knowledge. Therefore to be precise, a statement about the value of information on some uncertain variable must be accompanied by a statement about which decision variables (those variables under the control of the decision maker) may be adjusted in response to the impending information.

Specifically, suppose that important variables associated with a particular decision problem have been divided into uncertain state variables $(s_1,\ldots,s_K)$ and decision variables $(d_1,\ldots,d_M)$. We can, for example, calculate the value of information on state variable s_i given flexibility on decision variable d_j . This value is interpreted as the economic benefit gained by utilizing the information about the state variable i when setting the decision variable j . Since the combined cost of supplying the information and adjusting the decision variable may or may not outweigh this benefit, it may or may not be profitable to incorporate some piece of information into some aspect of the decision-making process. Being able to

calculate the value of supplying knowledge of various state variables while retaining flexibility of various decision variables should aid a manager in designing and timing information transmission for the purpose of decision-making.

A consideration almost as fundamental as the distribution of information is the accuracy with which it will be reported. A basic source of distortion in information transmission is quantization. Quantization takes place whenever physical quantities are represented numerically; further "rounding off" is likely to take place in the reduction and processing of data. Finally, because of the large number of variables in many decision problems, rough quantization of variables to as few as three levels is frequently necessary to reduce to a computation of manageable size.

We have found that we can make a calculation very similar to the value of information calculation that will allow us to place a dollar value on the distortion caused by a particular quantization scheme. Quantization in the information system can be thought of as a state space partitioning as follows. Rather than learning the exact value of a state variable prior to setting a decision variable, a quantized report tells us only that its exact value lies in one of a mutually exclusive, collectively exhaustive collection of subsets $S_1,\ldots,S_N$. Specification of these subsets is equivalent to specifying the quantizing scheme. Two important questions emerge: First, what is the value of a particular level of discretization? Second, how should the partition $S_1,\ldots,S_N$ be determined?

A thorough answer to these questions would lead to a means for determining and evaluating optimal quantized information systems for decision problems. So far our analysis has lead to several interesting results. The economic loss associated with the use of a particular number N of quantization levels tends to decrease rapidly with an increase in N . For example, if we have a uniformly distributed state variable and a quadratic value function, this cost decreases as $1/N^2$. Our preliminary results indicate that very rough quantization to as few as three levels gives relatively small distortion losses. We have also found that optimal quantizing schemes exist that give significantly better performance, and, in some cases at least, the optimal and quantizing partitions can be very easily constructed.

We are optimistic that continued research into the areas of flexibility and quantization in informational systems will continue to yield results of practical as well as theoretical importance to the decision maker. Just at a time when we are becoming more and more aware of the necessity for utilizing available information, the compilation and transmission costs of data are declining. Surely a framework is needed for balancing the economic benefits and costs associated with information usage.